Summary Report of the 46th Northeast Regional Stock Assessment Review Committee (SARC 46)

Members of SARC 46

Michael Murphy, chair
Chris Darby
Neil Klaer
Geoff Tingley

Prepared for NE Region Stock Assessment Workshop

National Marine Fisheries Service
National Oceanic and Atmospheric Administration
Woods Hole, Massachusetts

Meeting dates November 26-29, 2007
Report date January 2, 2008

Contents

Executive Summary

The 46th SARC (Stock Assessment Review Committee) met in Woods Hole, Massachusetts, from Monday, November 26, through Thursday, November 29, 2007, to review the assessment of Atlantic striped bass *Morone saxatilis*.

The review committee was composed of Mr. Michael Murphy (Florida Fish and Wildlife Conservation Commission, chair) and three scientists affiliated with the Center for Independent Experts, University of Miami: Dr. Chris Darby, Dr. Neil Klaer, and Dr. Geoff Tingley. The SARC was assisted by the NEFSC SAW Chairman, Dr. James Weinberg, his staff, and staff of the Atlantic States Marine Fisheries Commission.

The background information and assessment of striped bass was presented on behalf of the Atlantic States Marine Fisheries Commissions striped bass technical committee's stock assessment sub-committee (assessment team) by Mr. Doug Grout, Dr. Gary Nelson, and Ms. Beth Versak. The SARC requested some sensitivity analyses to the statistical catch-at-age analysis and additional background information.

The SARC concluded that the assessment team had successfully met all of their terms of reference. The extensive data available for the assessment appeared to be correctly compiled and used in the assessment and the analyses were made in accordance with good scientific practice.

The review committee found that, of the candidate assessment models, the statistical catch-at-age model (SCA) best estimated parameters that could be judged against the current biological benchmarks, 1995 spawning stock biomass and fully recruited fishing mortality rate at maximum sustainable yield. Based on these, the SARC agreed with the assessment team's stock status determination that striped bass were not currently overfished and overfishing was not occurring. Fishing mortality has increased in recent years and is currently (data up to and including 2006) at or very near the target level.

The review committee was impressed with spatially detailed data available for assessing striped bass. Suggestions were made that could reduce the difficulties encountered in the current global assessment model, e.g. conflicting indices, if the spatial distribution of the stock was considered in the assessment analyses. It was noted that this would make better use of the available data.

1 Introduction

1.1 Background

The 46th SARC (Stock Assessment Review Committee) met in Woods Hole from Monday, November 26, through Thursday, November 29, 2007, to review assessments of Atlantic striped bass *Morone saxatilis*.

The members of SARC were the chair, Mr. Michael Murphy of the Florida Fish and Wildlife Conservation Commission; and three scientists affiliated with the Center for Independent Experts, University of Miami: Dr. Chris Darby, Dr. Neil Klaer, and Dr. Geoff Tingley. The SARC was assisted by the NEFSC SAW Chairman, Dr. James Weinberg and by staff of the NMFS Northeast Fisheries Science Center and Atlantic States Marine Fisheries Commission.

About two weeks before the meeting, assessment documents and supporting materials were made available to the SARC via an ftp server. On the evening before the meeting, the assessment review committee met with Drs. James Weinberg and Paul Rago, NEFSC, to discuss the meeting agenda, reporting requirements, and meeting logistics. During the SARC meeting, all documents were available electronically and in print.

The meeting format included presentations mixed with questions and open discussion. The entire review committee participated in the review of each term of reference. The first 2-3 days of the meeting were open to the public and public comments were accepted during that time.

1.2 Review of Activities

The first day of the meeting (Monday afternoon) was devoted to presentations made by members of the ASMFC striped bass stock assessment subcommittee. Mr. Doug Grout presented background material on the ecology and biology of striped bass, the history of exploitation, management, and assessments, and a brief overview of the current assessment's findings. The review committee requested and received more information about male maturity at age and length at age.

Dr. Gary Nelson followed with detailed presentations on the information pertaining to the first three terms of reference. During discussions of TOR 1, characterization of commercial and recreational harvest and discards, the assessment review committee requested more information about historic levels of commercial harvest. Graphs were requested and made available showing that the commercial harvest, by weight had peaked in 1972 at nearly 6,000 mt (Fig. 1). This was about twice the 2006 commercial landings level. The 2006 recreational fishery harvest, 14,000 mt, was considerably higher than the peak commercial landings. It was noted by the assessment team that the sizes of striped bass harvested during the 1970's was much smaller than now so, in terms of numbers, historic landings were still considerable. Other discussions and questions pertained to the precision and bias of scale age determination, problems using the recreational discard and tag return rates to estimate commercial discards, sex ratio changes with size, historic changes in the commercial landings reporting, and catch-at-age validation through large cohort progressions.

Gary Nelson also presented the assessment team's characterization of the numerous indices of abundance developed for striped bass (TOR 2). Questions and discussion about the survey coverage, both temporal and spatial, were made. The final presentation made on the first day reported on the structure of a statistical catch-at-age model (SCA) developed for the analysis of striped bass. The review committee had questions about the choice and use of several selectivity models, and the choice of constant-selectivity periods. It was also noted that the model fit to the catch proportion-at-age (CAA) data for the overall harvest and dead discard was much tighter than the fit to the indices. The review committee requested a model run where the effective sample size for the CAA was set to about one-tenth of the current size to investigate a more even weighting scheme. The meeting adjourned at the NEFSC and the committee reformed at the hotel to discuss the day's presentations.

On Tuesday morning, the meeting reconvened and heard presentations of the findings derived from the Monday requests. The SCA re-run at lower weights on the CAA data, down to 5% of original effective sample size, showed little change in the predicted fully-recruited instantaneous fishing mortality rate (F) estimates though the precision of the estimates declined (Fig. 2). Additionally, there was little effect on the final estimates of F when abundance indices were also sequentially dropped from the input data.

For the rest of Tuesday morning, the review committee heard a presentation by Ms. Beth Versak and discussed Baranov's catch equation method (CEM, TOR 4) and a presentation by Doug Grout on biological reference points (TOR 7). The discussions involved exploration of potential sources of changes in the tagging programs, e.g., reporting rates and tag loss, that could account for the flat fishing mortality rates during a recent period of increasing catch and stable, but variable, recruitment. Before reviewing the biological reference point TOR 7, the review committee requested clarification of its intent. The assessment team responded that it saw its charge under this TOR as simply to determine the stock status of striped bass but that the review committee should comment on the current biological reference points. The review committee requested an updated spawner-recruit plot (Fig. 3).

The review committee heard Beth Versak's presentations on the instantaneous rates tag return model (IRCR, TOR 5) and Gary Nelson's presentation on the statistical catch-at-age with tag data model (SCATAG, TOR 6) on Tuesday afternoon. In general, the committee was concerned about how the estimated annual fishing mortality rates were determined using model averaged estimates that had contributions from period models (where a single fishing mortality was estimated for a period of years). The review committee found that the SCATAG model appeared a reasonable approach but was concerned about how it (and the SCA) ignored the geographic structuring of the striped bass stocks. Finally, the review committee met alone for a short time to discuss its findings for each TOR. The stock assessment committee was then briefly presented with these findings and invited to discuss them with the committee during the final hours of the meeting on Tuesday.

Report writing began in earnest on Wednesday when the review committee drafted sections of the summary report. Requests for further information and graphs were made of the assessment team during the morning and afternoon. Writing continued on Thursday morning until the meeting adjourned at about noon.

2 Review of striped bass assessment

2.1 Terms of reference

The SARC considered the striped bass assessment in light of the terms of reference (TOR) provided to the SAW, as follows:

1. Characterize the commercial and recreational catch including landings and discards.
2. Characterize the fisheries independent and dependent indices of abundance.
3. Evaluate the Statistical Catch at Age (SCA) model and its estimates of F, spawning stock biomass, and total abundance of Atlantic striped bass, along with the uncertainty of those estimates.
4. Evaluate the Baranov's catch equation method and associated model components applied to the Atlantic striped bass tagging data. Evaluate estimates of F and abundance from coast-wide and Chesapeake Bay specific programs along with the uncertainty of those estimates.
5. Review the Instantaneous Rates Tag Return Model Incorporating Catch-Release Data (IRCR) and estimates of F on Atlantic striped bass. Provide suggestions for further development of this model for future use in striped bass stock assessments.
6. Review the Forward-Projecting Statistical Catch-At-Age Model Incorporating the Age-Independent Instantaneous Rates Tag Return Model (SCATAG) and estimates of F, spawning stock biomass, and total abundance of striped bass. Provide suggestions for further development of this model for future use in striped bass stock assessments.
7. Evaluate the current biological reference points for Atlantic striped bass from Amendment 6 and determine stock status based on those reference points.

2.2 SARC findings by term of reference

2.2.1 TOR1 Commercial and recreational catch characterization

The review committee (SARC) found this term of reference was met. The complexity of the development of the commercial and recreational harvest and discard estimates were described well in the assessment report documentation. The review committee found that the commercial harvest information was based on a complex and changing series of collection programs that probably recorded most of the harvest. The commercial discards had to be estimated by projecting the commercial-to-recreational tag discard ratio onto the estimated recreational discards. While this may be adequate, the review committee suggests that the associated error for this estimate be included in the data characterization and carried forward in the assessment. The recreational harvest and discard information was well characterized but the committee found that the apportioning of discards into those that survived and those that died should reflect the published rates of hooking mortality and their seasonal effects. Currently, a single overall release mortality proportion of 0.08 is used for historic consistency. Published information (not seen by the review committee) now puts the average rate at 0.09 and suggests some seasonal differences that should be used in estimating dead discards. This is important because even at the currently estimated levels, the discarded dead is a significant portion of the harvest and small changes to this component could be important in the assessment.

The assessment team reported that age determination for striped bass older than about age 10 is biased using scales, based on comparisons made for a small number of samples with concurrent otolith collections. This problem needs to be remedied through the collection of more otolith samples from older fish. It is especially important to characterize the selectivity of these older fish because selectivity may decline with age and stock status is partially judged against the level of F for older fish.

The unaccounted mortalities from a potential high-grading process or unreported commercial discards in the EEZ are apparently minor but periodic monitoring of these potential sources of mortality should be carried out.

The review committee was briefly apprised of the current status of the NMFS Marine Recreational Fisheries Statistics Survey. Though there may be some bias in this survey with regards to striped bass fishing, e.g. lack of coverage in freshwater areas of estuary, these data are the best available for the recreational fishery. In fact, recent coverage for striped bass fishing is enhanced because the ASMFC requires add-on intercepts in coastal states where the proportional standard error for the estimated harvest is greater than 20%.

2.2.2 TOR2 Fishery-dependent and fishery-independent indices

The SARC concluded that the fisheries independent and dependent indices of abundance had been well characterized and that this term of reference was met. A diverse range of indices is available, provided by various Federal and State agencies independently over considerable time periods. A number of technical reviews in recent years have examined the indices used for stock assessment leading to revisions and elimination of some indices (ASMFC (1996) review of young-of-the-year indices and 2004 workshop review of indices)

Fisheries dependent catch rate indices were Massachusetts Commercial Total Rate Index (MACOMM), Connecticut Recreational CPUE (CTCPUE) and MRFSS Total Catch Rate Index (MRFSS). Fisheries independent indices for combined ages 2-13+ were the Connecticut Trawl Survey (CTTRL), Northeast Science Center Bottom Trawl Survey (NEFSC), New Jersey Bottom Trawl Survey (NJTRL), New York Ocean Haul Seine Survey (NYOHS), Maryland Spawning Stock Survey (MDSSN) and Delaware Spawning Stock Electrofishing Survey (DESSN). Fishery independent young of the year and age 1 indices of abundance were the New York Young of the Year (YOY), New York Age 1, New Jersey YOY, Virginia YOY, Maryland YOY and Maryland Age 1.

The assessment team explained some additional modifications to indices including discarding data for the NEFSC survey prior to 1991 due to design changes and lack of some inshore strata in the earlier years.

There is some agreement among the fishery-dependent indices in that they all show an increasing trend. However the patterns differ, with the MACOMM and MRFSS being on average flat from about 2000 onwards and CTCPUE showing an increase over that period. It was noted that there is some redundancy in the MRFSS and CTCPUE indices as these both used the same effort series.

Fisheries-independent aggregated (age $2\text{-}13^{+}$) indices show differing trends. The CTTRL from 1984 and NJTRL from 1990 show a generally increasing trend, perhaps flattening in recent years. The NEFSC, MDSSN and DESSN are generally flat for the periods they cover, and the NYOHS indicates some increase from 1986 to 2006 but is highly variable. Inter-annual variability in all of these indices is high.

The young of year indices probably provide more information than is used in the current assessments. They provide direct information about differences in recruitment levels for the sub-stocks of the fishery and differences in pattern by region are clear.

In previous years, differing methods of averaging have been used for different indices – either arithmetic or geometric mean. In this year's assessment presented to the SARC, the stock assessment team has attempted to standardize the averaging and chose arithmetic averaging. The committee notes that generalized linear modeling (GLM) procedures, for example, usually assume a log-normal error structure that applies to abundance indices, and suggests that a geometric mean be used to average indices prior to input to stock assessment.

There are differences in methodologies applied by the various organizations in how surveys are carried out – especially in random stratification. It is recognized that time series should be maintained where possible, but there also appears to be scope for further standardization of procedures used to conduct the surveys among organizations. Standardization of procedures would also better facilitate the development of indices that can be collected within state regions but are also comparable enough that they could be combined over larger areas. There is a need for the collection of more fishery-wide fishery-independent index data and for more regional (State) CPUE data from the recreational fisheries.

The amount of information provided by these indices on abundance trends of different age classes or age groupings in the striped bass population in different regions is high compared to most fisheries, and the stock assessment team has incorporated them all within the stock assessment. However there is still scope for rationalization, either by further combination of indices before introduction into the assessment, or better accounting for differences among indices within the assessment, particularly by region.

As many of the indices are regional in nature, the review committee feels that there is scope to incorporate spatial differences in index analyses and averaging before use in the assessment. Alternatively, the assessment needs to include spatial structure to make better use of index information.

The committee notes that GLM procedures used to standardize catch rate indices differ and that there is scope to apply standard methods.

2.2.3 TOR3 Evaluate the statistical catch at age (SCA) model

The assessment team met this term of reference. The review committee found that it was the best model for determining stock status at this time. The assessment team is commended for its clear presentations covering the fitting of the models and the investigations that had been carried out to explore and develop alternative model structures to those adopted for previous assessments. Interim modeling and analytical studies have clearly progressed and improved the modeling of the dynamics of the striped bass stocks and the SCA model is considered to be an improvement over the previously used ADAPT framework.

The model is fitted within an appropriate statistical framework and the assessors have a clear understanding of the fitting procedures and the interpretation of the model output.

The document describing the model results provides details of the analysis and the uncertainties for review and the assessment team's interpretations of the model fit and estimated trends in the stock metrics are appropriate.

The review committee considered that there were several areas which should be examined by the assessment team in future developments of the model, which fall into two categories:

(1) Development of the current single stock hypothesis model
(2) Development of a spatially segregated multi-stock model

(1) Development of the current single stock hypothesis model

The SCA presented is appropriately fitted to the combined catch proportion at age (CAA) data set raised from the commercial and recreational landings and discards described in TOR 1. However, the model fit is dominated by the fit to the CAA data set at the true ages (it is especially poor at modeling the dynamics of the plus group for all data sets). Sensitivity analysis showed that the model captures the overall trends in many of the survey series but it does not capture the dynamics of individual ages through time. The committee noted that the effective sample size given to the CAA data could be resulting in a weighting of the catch at age data that was too high and requested a reduction to 1/10th of the value used within the base analysis. The resulting time series of mortality estimates are almost identical to those estimated by the heavily weighted model but the uncertainty associated with them is increased (Figure 2).

The review committee considers that the models inability to capture the dynamics of any of the survey series is the result of the difference in the spatial scale of the catch based assessment and the localized area over which the surveys are collected. The surveys reflect changes in the local dynamics of the three sub-stocks and without analysis to combine the time series into an aggregated stock index (see comments on TOR 2 and (2) below) there will always be an incompatibility that introduces uncertainty into the model estimates. If a single stock model is to be continued a spatially aggregated index for each age or groups of ages (e.g. spawning biomass) should be developed for use as a calibration series for fitting the assessment model.

Model mis-specification is most obvious at age 13^+. The plus group is not modeled appropriately for the catch at age or for any of the survey indices. This would suggest that a re-specification of the dynamics of the plus group is required. One possibility that should be explored is use of a dome shaped selectivity function.

The review committee considered that whilst the SCA model is closely fitting the aggregated catch proportion at age data set this may not be the most appropriate way to model the individual components contributing to stock mortality (recreational landings, recreational discards, commercial landings and commercial discards) and to capture the uncertainty associated with measuring each component. Consideration should be given to modeling the individual data components separately in order to capture the dynamics and uncertainty associated with each data set rather than aggregating and smoothing through the uncertainty.

Ideally the modeling of the separate components should include the tagging data within a combined model similar to the SCATAG model presented in TOR 6.

Whilst the assessment report describing the fit of the SCA model to the available data is sufficient for determining that the model is fitted appropriately to the catch proportion at age data, the review committee considered that insufficient diagnostic output had been presented. There are a number of statistical methods useful for analysis of residual patterns and model-fit diagnostic graphical plots that are commonly used for evaluating stock assessment model fit (for example q-q plots). The committee considered that the assessment team should spend time developing diagnostic output tables and figures that would enhance their ability to examine alternative model fits and runs. Residual plots should describe the type of residuals being used for presentation and should preferably be standardized. The standard deviation of the standardized residuals provides a good measure of the relative model fit to the data among datasets.

The inclusion of stock and recruitment model estimation within the model fit, given the fairly well defined relationship estimated in the current model, would provide a useful routine check on the appropriateness of the management reference points and would allow stock projections if required should mortality exceeds thresholds.

(2) Development of a spatially segregated multi-stock model

The amount of information on a variety of sources and life history stages that is available for the analysis of the dynamics of the striped bass stock is considerable, providing one of the most data rich environments that the reviewers have seen. The information is available to the analysts as time series that are spatially and temporally disaggregated and it is considered that the current assessment process (SCA and tagging models) is not making full use of the information that is available.

The use of the current single stock hypothesis model when there is clear evidence for three sub-stocks with differing contributions to the offshore mix of adults and apparently (from local survey trends) differing dynamics, argues for a more complex, spatially separated assessment model. Such assessment models require information on the dynamics of each of the components, an ability to separate catches by region and perhaps migration rates between regions (tagging data). All of this information is available for striped bass.

Although the assessment is based on a unit stock, the management of striped bass is implemented at a regional level. Therefore, in order to provide management advice at that level it would be advantageous to managers to be able to determine the relative dynamics of the sub-stocks. Assessment models that are fitted to multiple stock units and which include mixing have been described in the literature. In general they suffer from a scarcity of migration and mixing (within catch by region) information; this is available for striped bass from the tagging studies. Such models would be ideally suited to the provision of more discrete scientific advice on stock dynamics to managers.

2.2.4 TOR4 Evaluate Baranov's catch equation method (CEM) for tagging

The review committee felt that this term of reference was met; however, the committee had reservations about the validity of the estimated fishing mortality and therefore used the SCA model output to compare to the biological reference point values.

The Baranov's catch equation method (CEM) applied to tagging data was presented as an alternative estimator of current fully recruited instantaneous fishing mortality (F) to the SCA model. The methodology is incorporated into the MARK program as used in previous years, including bias adjustment to annual estimates of survival. Previously a fixed value of the instantaneous natural mortality (M) was used to derive F estimates from annual survival rates produced by MARK. This year, following methods provided by Ricker (1975), an annual exploitation rate was estimated using the number of fish tagged at the beginning of the year, those recaptured and killed and those recaptured and released alive. Assumptions included a released survival rate of 0.92 and a reporting rate of 0.43. M was then derived by subtracting estimated F from estimated instantaneous total mortality (Z). Stock size was estimated using the ratio of average stock size and F. These calculations were applied to ≥ 18 inch and ≥ 28 inch fish size groups, four coastal programs and four producer area programs.

There were nine different candidate models used in the MARK program encompassing survival and reporting that was held constant for all years (constant), within regulatory period or was estimated annually (time specific). Annual survival values from the MARK program were produced by averaging results over all of these models based on Akaike weights assigned to each model.

Coastal and producer area results were calculated by averaging across area-specific results using a proportional contribution weighting. The proportional contribution by area is from work completed in the 1970's. In Chesapeake Bay the split between MD and VA is based on the area of the spawning grounds.

The coast-wide fishing mortality values were produced from the arithmetic averages of the coastal and producer area overall results.

The Committee noted that the assumption of 0.92 release survival is inconsistent with the published value of 0.91 (see TOR 1), and agrees with the assessment team that the reporting rate is unlikely to be constant.

The Committee questioned whether the stock size calculated here is really the vulnerable stock size, and whether it was comparable to stock size from SCA. There is an assumption that everything caught and tagged is fully recruited.

There may be an underestimation of hooking-mortality F because fish caught and released and not reported are assumed to survive.

The assessment team was questioned about whether it was appropriate to apply a constant reporting rate of 0.43 from the recreational fishery to the commercial fishery. The team agreed that the reporting rate from commercial fisheries is probably lower than for the recreational fishery. However, there is some evidence from the MD high reward tagging study that the reporting rate could be higher than the estimate used for both the recreational and commercial fisheries. If the reporting rate is underestimated, the resulting F values are overestimated, so the current procedure is conservative from that viewpoint. The assessment team described current additional work that should determine whether there is a difference in the coast-wide and bay reporting rates. These data should become available in 2008.

The review committee noted that the 36th SAW advisory committee had recommended the removal of the constant survival tagging model as it was not biologically reasonable given documented changes in fishing effort. The review committee agrees with this earlier recommendation.

There was concern about negative M estimates for certain tagging programs and how these might be included in area averages. The assessment team is currently not including negative values in the averages.

Results indicate increasing M in coastal regions outside Chesapeake Bay where mycobacteriosis occurs and may be causing increased mortality. The assessment team believes that this may be due to fish moving out of the Bay into coastal waters. Maryland Department of Natural Resources monitors prevalence of the disease which may currently be as high as 75% of fish in the Bay. However, there are no monitoring and no earlier estimates to determine whether the prevalence has been increasing. In general, the assessment team has not demonstrated that the missing fish that are being assigned to M are not lost to other causes of "unaccounted removals" such as changes in migration rates and patterns or tag reporting changes.

The review committee noted that stable recruitment and increasing catches over recent years would normally mean that F has been increasing. The recent trends in estimated F's using Baranov's catch equation method (CEM) have been flat, with an increasing portion of total mortality shifted to an increase in M. Methods for detection of increased natural mortality should be investigated – either field or model studies. Changes in the reporting rate over time should also be investigated as a possible source of false changes in the survival rate.

The review committee requested a comparison of SCA estimates of stock size for 7+ with those from Baranov's catch equation method. The estimates were of a similar magnitude. SCA estimates were higher from 1990-2002, and CEM estimates higher from 2003-2006.

The review committee was concerned about the very high estimates of M for MD and VA for the Chesapeake Bay specific analysis. Many values are greater than 0.6 yr^{-1} in MD and over 1.0 yr^{-1} in VA which are not biologically plausible.

2.2.5 TOR 5 Review the Instantaneous rates tag return model (IRCR)

The review committee found that this term of reference was met.

This IRCR model offers an additional approach to using the substantial tag return data to assess the striped bass population. The model allows for released fish to be directly incorporated into the estimation of F, with no *ad-hoc* bias-correction necessary. It follows the age-independent approach presented in Jiang *et al.* (2007) and has two input matrices: one contains information only about harvested fish and the other contains only information about released fish that had their tags cut off.

The observed recovery matrices are compared to expected recovery matrices to estimate model parameters and the expected values follow a multinomial distribution so the full likelihood is the product multinomial of the cells. The model was programmed in AD Model Builder.

In considering this approach, the review committee noted that there were shared issues with this and the other tag based models. Particularly, this includes the unpredictable selection of the various sub-models that define the results.

Observations by the review committee included the IRCR estimation of M as 0.15, but the output F is substantially different from the SCA model. The cause of this difference remains unresolved.

The presentation of confidence intervals around the coast-wide estimates would have been helpful in interpretation.

It was noted that estimates of M from the IRCR differ between the different areas. Although this was not understood it is possible that this was derived from some aspects of the logistics of the tagging programs – differences in the size of fish tagged, time of year tagged, or in environmental aspects (e.g. temperature) of tagging areas, etc.

The committee noted that this approach did have a number of advantages over the CEM. This model does address one of the problems of the catch-equation method, where the data are used in two ways where there was a concern that data were not handled in a consistent manner but then used together.

One benefit of this model is that it constrains the output so that the negative values of F or M seen in runs of previous models do not occur. An added benefit of this model is the ability to include released fish in the analysis.

The assessment team had not yet tried approaches other than simple model averaging, which might yield improved outputs.

A serious concern identified was the added complexity of this model, as there were three products: F, F' (mortality on tags), and M.

There was concern about the validity of averaging F's across program-specific studies. The assessment team has looked at this but there appeared to be a mismatch between the various data collection programs from the timing of the tagging and recapture year, and it seems that whichever program has the largest sample size, influences the outcome the most, which is undesirable.

To conclude, the review committee considered that this was a fruitful approach, but was possibly overestimating the information content of the data.

2.2.6 TOR 6 Review the **SCA with tagging data (SCATAG)**

The review committee received a presentation from the assessment team on the progress made towards developing an integrated catch at age and tagging analysis assessment model.

The review committee found that this term of reference was met.

The first attempt at a combination model, SCATAG, is currently at an early stage of development. It links the age-based SCA model (described in TOR 3) and the age independent IRCR tagging analysis (described in TOR 5) through a single time series of estimated fully-recruited year effects in fishing mortality. The model assumes constant selection at age within a year in the SCA model at the oldest ages, and constant selection at length in the tagging experiments for fish greater than 28 cm.

Initial model fits were presented to illustrate that the model converged and reproduced historic trends in fishing mortality and stock abundance that were consistent with the individual model estimates. The more recent stock and mortality dynamics were consistent with the IRCR time series. Some sensitivity analyses had been examined using changes to the component model weights in the likelihood function but these were at an early stage of development and considered to be model exploration exercises.

The committee welcomed the development time that had been assigned to the model and considered that the combination of a catch and tag based model was a worthwhile progression for this stock; discussions with the assessment team covered the developments to date and the initial model results and sensitivity.

There was concern that the two model components are estimating differing fishing mortality metrics that are not directly comparable. Therefore, assuming that the fishing mortality values are directly comparable across the models may be introducing over or under estimation bias in the combined series of fishing mortality estimates. A more consistent approach used in other studies is to apply an age length matrix to the length distribution of the tagged fish in order to standardize the fishing mortality year effects between models.

It has been noted previously that the tagging model estimates uses a constant selection at length for large fish and the age-based model constant selection at age at the oldest ages. The possibility of dome shaped selection in catchability with increasing size was discussed and it was noted that studies in Canada had successfully used tagging data to estimate selection curves and that similar analysis could be applied to the striped bass tag returns.

Concern was raised about the lack of sensitivity in the model estimates to the down-weighting of likelihood components and the potential for correlation in the information used to fit the model. As noted in the discussions on the SCA model (TOR 3) the catch proportion at age matrix dominates the model fit and gives very little weight to the survey time series. Adding the tagging-data model induces time series of estimates of fishing mortality and stock abundance that mimic the trends from the tagging model in isolation. This strongly suggests that there is very limited information contributed to the likelihood from the fit to the catch proportion at age and the survey information is effectively ignored within the model fit. This once again raises questions regarding the weighting of the components within the likelihood.

2.2.7 TOR 7 Evaluate the current biological reference points.

The meaning of this term of reference was clarified just prior to the presentation of the TOR by the assessment team. In addition to determining the stock status, its purpose was to review the methods used to determine the current biological reference points and to get the review committee’s opinion on whether they were developed appropriately and whether those approaches should be continued. Based on this clarification, this term of reference was successfully completed. Fishing mortality has increased in recent years and is currently (data up to and including 2006) at or very near the target level.

The committee noted that the reference points for this species are management oriented rather than biologically based and that these have been effective in promoting stock recovery. There should be a link between the threshold and target, such that if the threshold is changed the target is also re-evaluated.

The reference points were estimated using an assumed sex ratio in the stock of 1:1 at all ages. Evidence was not presented to support a 1:1 sex ratio from the catches. Given the differential rates of growth, and differences in spatial and temporal occurrence (migration and distribution) of the two sexes, this should be reviewed to ensure that the effect of this assumption is negligible. Alternatively, means to better estimate the sex ratio by age in the population should be implemented.

The fishing mortality at maximum sustainable yield (F_{MSY}) was estimated using the outputs of a virtual population analysis (VPA) and yield per recruit analysis. Ideally this needs to be both current and compatible with the current assessment methodology. The implication of this is that the value of F_{MSY} should be recalculated using a model of the same type as that of the preferred assessment methodology. This is especially important given the high numbers of discards of young fish in this fishery because the calculation of F_{MSY} is affected by the selectivity pattern not just the F on fully recruited fish.

The tagging (length-based) models are measuring an F that is not directly related to the F generated from the SCA model. Based on the evidence presented, the review committee believes that SCA-generated F is better suited to provide a comparator for reference points. The F estimated from tag/recapture data is an aggregate fishing mortality rate (catch=F*[average abundance]) estimated for the sample of tagged fish. In as much as this sample is representative of the population, it should be approximately equal to the aggregate rate that is estimable from the SCA-based estimates of average abundance and catch. Since the aggregate rate is affected by changes in selectivity, it is not necessarily comparable to the fully recruited F stipulated as the metric to be compared with the reference-point F_{MSY} unless only fully recruited fish are included in the tagging sample. This was approximated using only tagged fish that were greater than or equal to 28 inches in the CEM analysis. The review committee was concerned that not all striped bass of these sizes are fully recruited, i.e., selectivity for striped bass may not be flat-topped. Other issues with the CEM method, e.g., changes in reporting rate, mortality of released-and-unreported tagged fish, the effect of averaging candidate models that estimate F /yr for different-length time periods (see TOR 4), added to the review committee's uncertainty about using CEM-based F's for determining striped bass stock status.

Within the current approach to the definition of reference points is the use of an absolute biomass estimate for 1995, which represents the ADAPT VPA estimate for 1995 biomass (deemed the year of striped bass recovery as generated by a forward projecting model of age-0 recruits[1]). Developments in stock assessments will mean that the absolute 1995 estimate from current and future stock assessments will vary from year to year. It therefore makes sense that the biomass reference point be measured relative to the most current estimate of 1995 biomass, e.g., $SSB_{threshold}$ = 1.1 estimated SSB_{1995}.

There is a known issue of potential bias in the aging of older fish using scales, a problem that has been considered at meetings and a specific technical workshop. The current reference points incorporate this bias as they were estimated using age-based methods using all age classes. This bias can be removed from the reference points by using a constrained age approach that only uses those ages where aging is know to be problem free. As this would incorporate ages of 10 and less, this would still utilize the majority of the available data and would be robust.

2.3 SARC recommendations on striped bass

1. The error in the catch estimation should be carried into the assessment, particularly for discards (TOR 1).
2. A workshop should be convened to compare age determination between scales and otoliths that are currently on hand (TOR 1).
3. Well-supported changes to key parameters within the assessment models should be used to update the model, such as release mortality rates and tag reporting rates (TOR 1).
4. There is scope to either further rationalize or combine indices prior to use in the assessment or to deal with them more specifically within the assessment. In particular, spatial information in the surveys should be better captured and accounted for prior to inclusion in a one-stock assessment. Alternatively, the assessment needs to account for spatial differences in the indices (TOR 2).
5. There is scope for further standardization of survey methodologies and also index standardization procedures employed by various organizations. Methods should be developed to standardize current sampling procedures or to develop new programs for the collection of fishery-wide fishery-independent index data (TOR 2).
6. Establish a set of regional fishery stock status indicators that can be examined independently from the assessment model that provide a spatial description of the health of the stock (TOR 2).
7. The SCA model should consider separate fits to the landings and discard data for the recreational and commercial fleets (TOR 3).
8. The SCA model framework should be expanded to account for the spatial dynamics of the striped bass complex and its fishery (TOR 3). The model should include fitting of the tag data within its spatial structure also. The spatial aspect of all of these recommendations needs to be considered.
9. The SCA model fit should be explored using better diagnostics.

---

1. ASMFC 1998. Amendment #5 to the Interstate Fishery Management Plan for Atlantic Striped Bass. Washington (DC): ASMFC. Fisheries Management Report No. 24. 31 p.

10. The stock-recruit relation should be incorporated into the SCA (TOR 3) and used to provide the F_{MSY} biological reference point. This relation and the reference point should be re-estimated each time the assessment is updated. The 1995 biomass reference point is appropriate if used as a relative measure rather than an absolute (TOR 7).

11. Develop field or modeling studies that will aid in estimation of natural mortality or other factors that affect the return rates of tags (TOR 4).

12. The model-averaged approach to estimating annual F's needs to be checked for its validity and sensitivity to year groupings (TOR 4).

13. Different models for selectivity within the plus group should be explored (including dome-shaped (TOR 4)).

14. Comparisons should not be made between stock metrics derived from tagging models and reference points derived from catch-at-age analysis unless differences in their derivation are accounted for (TOR 7). Ideally, reference points should be model-specific.

3 Acknowledgements

The review committee received valuable assistance from rapporteurs Gary Shepherd, NEFSC, and Nichola Meserve, ASMFC. Discussions during the meeting benefited from the participation of NEFSC staff Gary Shepherd, Paul Rago, Chris Legault, and Elizabeth Brooks. The NEFSC SAW/SARC staff also helped facilitate the conduct of the meeting and completion of this report.

References

ASMFC. 1996. Report of the juvenile abundance indices workshop. Washington, D.C.: ASMFC Special Report No. 48 83p.

Jiang, H., K.H. Pollock, C. Brownie, J.M. Hoenig, R.J. Latour, B.K. Wells, and J.E. Hightower. 2007. Tag return models allowing for harvest and catch and release: evidence of environmental and management impacts of striped bass fishing and natural mortality rates. N. Amer. J. Fish. Manage. 27:387-396.

Appendix A Statement of Work

SARC 46: Statement of Work for CIE Reviewers (including a description of SARC Chairman's duties)

General

The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development (SAW Working Groups or ASMFC technical committees), assessment peer review, public presentations, and document publication.

The SARC46 review panel will be composed of three appointed reviewers from the Center of Independent Experts (CIE), and an independent chair from the Florida Fish and Wildlife Conservation Commission. The panel will convene at the Woods Hole Laboratory of the Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts, from November 26 - 29, 2007 to review one assessment (Striped bass, *Morone saxatilis*). In the days following the review of the assessments, the panel will write the SARC Summary Report and each CIE reviewer will write an individual independent review report.

Specific Activities and Responsibilities

The CIE's deliverables shall be provided according to the schedule of milestones listed on Page 5. The CIE reviewers, along with input from the SARC Chairman, will write the SARC Summary Report. In addition, each CIE reviewer will write an individual independent review report. These reports will provide peer-review information for a presentation to be made by NOAA Fisheries at meetings of the New England and Mid-Atlantic Fishery Management Councils in 2008. The SARC Summary Report shall be an accurate and fair representation of the SARC panel viewpoint on how well each SAW Term of Reference was completed (please refer to Annex 1 for the SAW Terms of Reference).

The three SARC CIE reviewers' duties shall occupy a maximum of 14 days per person (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; and several days following the open meeting to contribute to the SARC Summary Report and to produce the Independent CIE Reports).

Not covered by the CIE, the SARC chair's duties shall occupy a maximum of 15 days (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; several days following the open meeting for SARC Summary Report preparation.)

Charge to SARC panel

The panel is to determine and write down whether each Term of Reference of the SAW (see Annex 1) was or was not completed successfully during the SARC meeting. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. Where possible, the chair shall identify or facilitate agreement among the reviewers for each Term of Reference of the SAW.

If the panel rejects any of the current Biological Reference Point (BRP) proxies for B_{MSY} and F_{MSY}, the panel should explain why those particular proxies are not suitable and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs are the best available at this time.

Roles and responsibilities

(1) Prior to the meeting

(SARC chair and CIE reviewers)
Review the reports produced by the Working Groups and read background reports.

(2) During the Open meeting

(SARC chair)
Act as chairperson, where duties include control of the meeting, coordination of presentations and discussion, making sure all Terms of Reference of the SAW are reviewed, control of document flow, and facilitation of discussion. For each assessment, review both the Assessment Report and the Assessment Summary Report.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to discuss the stock assessment and to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

(SARC CIE reviewers)
For each stock assessment, participate as a peer reviewer in panel discussions on assessment validity, results, recommendations, and conclusions. From a

reviewer's point of view, determine whether each Term of Reference of the SAW was completed successfully. Terms of Reference that are completed successfully are likely to serve as a basis for providing scientific advice to management. If a reviewer considers any existing Biological Reference Point proxy to be inappropriate, the reviewer should try to recommend an alternative, should one exist.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

(3) After the Open meeting

(SARC CIE reviewers)
Each reviewer shall prepare an Independent CIE Report (see Annex 2). This report should explain whether each Term of Reference of the SAW was or was not completed successfully during the SARC meeting, using the criteria specified above in the "Charge to SARC panel" statement.

If any existing Biological Reference Point (BRP) proxies are considered inappropriate, the Independent CIE Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.

During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments may be raised. Comments on these questions should be included in a separate section at the end of the Independent CIE Report produced by each reviewer.

If a reviewer feels that his/her comments are adequately expressed in the SARC Summary Report, it will not be necessary to repeat the same comments in the Independent CIE Report. In that case, the Independent CIE Report can be used to provide greater detail on specific Terms of Reference or on additional questions raised during the meeting.

(SARC chair)
The SARC chair shall prepare a document summarizing the background of the work to be conducted as part of the SARC process and summarizing whether the process was adequate to complete the Terms of Reference of the SAW. If appropriate, the chair will include suggestions on how to improve the process. This document will constitute the introduction to the SARC Summary Report.

(SARC chair and CIE reviewers)
The SARC Chair and CIE reviewers will prepare the SARC Summary Report. Each CIE reviewer and the chair will discuss whether they hold similar views on each Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the SAW. For terms where a similar or a consensual view can be reached, the SARC Summary Report will contain a summary of such opinions. In cases where multiple and/or differing views exist on a given Term of Reference, the SARC Summary Report will note that there is no agreement and will specify - in a summary manner – what the different opinions are and the reason(s) for the difference in opinions.

The chair's objective during this Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement if it cannot reach one. The chair will take the lead in editing and completing this report. The chair may express the chair's opinion on each Term of Reference of the SAW, either as part of the group opinion, or as a separate minority opinion.

The SARC Summary Report (please see Annex 3 for information on contents) should address whether each Term of Reference of the SAW was completed successfully. For each Term of Reference, this report should state why that Term of Reference was or was not completed successfully. The Report should also include recommendations that might improve future assessments.

If any existing Biological Reference Point (BRP) proxies are considered inappropriate, the SARC Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRP proxies are the best available at this time.

The contents of the draft SARC Summary Report will be approved by the CIE reviewers by the end of the SARC Summary Report development process. The SARC chair will complete all final editorial and formatting changes prior to approval of the contents of the draft SARC Summary Report by the CIE reviewers. The SARC chair will then submit the approved SARC Summary Report to the NEFSC contact (i.e., SAW Chairman).

Schedule

The milestones and schedule are summarized in the table below. No later than December 17, 2007, the CIE reviewers shall submit their Independent CIE Reports to the CIE Program manager Dr. Manoj Shivlani via e-mail to mshivlani@rsmas.miami.edu

Milestone	Date
Open workshop at Northeast Fisheries Science Center (NEFSC) (begin writing reports, as soon as open Workshop ends)	November 26-28, 2007
SARC Chair and CIE reviewers work at the NEFSC drafting reports	November 28-29
Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair **	December 17
CIE reviewers submit Independent CIE Reports to CIE for approval	December 17
SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)	December 24
CIE provides reviewed Independent CIE Reports to NMFS COTR for approval	December 31
COTR notifies CIE of approval of reviewed Independent CIE Reports	January 7, 2008 *
COTR provides final Independent CIE Reports to NEFSC contact	January 7, 2008

* Assuming no revisions are required of the reports.

** The SARC Summary Report will not be submitted, reviewed, or approved by the CIE.

The SAW Chairman will assist the SARC chair prior to, during, and after the meeting in ensuring that documents are distributed in a timely fashion.

NEFSC staff and the SAW Chairman will make the final SARC Summary Report available to the public. Staff and the SAW Chairman will also be responsible for production and publication of the collective Working Group papers, which will serve as a SAW Assessment Report.

NEFSC Contact person and SAW Chairman:
Dr. James R. Weinberg, NEFSC, Woods Hole, MA. 508-495-2352, James.Weinberg@noaa.gov

Submission and Acceptance of CIE Reports

No later than December 31, 2007, the CIE shall provide via e-mail the final independent CIE reports and the CIE chair's summary report to the COTR William Michaels (William.Michaels@noaa.gov) at NOAA Fisheries. The COTR and alternate COTR Dr. Stephen K. Brown (Stephen.K.Brown@noaa.gov) will review the CIE reports to determine that the Term of Reference was met, notify the CIE program manager via e-mail regarding acceptance of the reports by January 7, 2008, and then distribute the reports to the NEFSC contact person.

Appendix B.

ANNEX 1:
Draft Assessment Terms of Reference
for the 46th Northeast Regional Stock Assessment Workshop

NOTE:
TORs listed below are preliminary. They were drafted in January, 2007. Final TORS will be available in September 2007, and will be provided to NOAA and the CIE.

A. Atlantic striped bass

1. Characterize the commercial and recreational catch including landings and discards.

2. Characterize the fisheries independent and dependent indices of abundance.

3. Review the catch at age based model used in the stock assessment to provide estimates of F, spawning stock biomass and total abundance and characterize the uncertainty of those estimates.

4. Review the tag based model used in the stock assessment to provide estimates of F and total abundance and characterize the uncertainty of those estimates.

5. Evaluate the biological reference points for striped bass and determine stock status based on those reference points.

ANNEX 2: Contents of SARC CIE Independent Reports

1. For each assessment reviewed, the report should address whether each Term of Reference of the SAW was completed successfully. For each Term of Reference, state why that Term of Reference was or was not completed successfully. To make this determination, CIE reviewers should consider whether the work provides a scientifically credible basis for developing fishery management advice. Scientific criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable.

   The report may include recommendations on how to improve future assessments.

   If a reviewer feels that his/her comments are adequately expressed in the SARC Summary Report, it will not be necessary to repeat the same comments in the Independent CIE Report. In that case, the Independent CIE Report can be used to provide greater detail on specific Terms of Reference or additional questions raised during the meeting.

2. If any existing Biological Reference Point (BRPs) proxies are considered inappropriate, include recommendations and justification for alternative proxies. If such alternatives cannot be identified, then indicate that the existing BRPs are the best available at this time.

3. Any independent analyses conducted by the CIE reviewers as part of their responsibilities under this agreement should be incorporated into their Independent CIE Reports. It would also be helpful if the details of those analyses (e.g, computer programs, spreadsheets etc.) were made available to the respective assessment scientists.

4. Additional questions that were not in the Terms of Reference but that are directly related to the assessments. This section should only be included if additional questions were raised during the SARC meeting.

ANNEX 3: Contents of SARC Summary Report

1. The main body of the report shall consist of an introduction prepared by the SARC chair that will include the background, a review of activities and comments on the appropriateness of the process in reaching the goals of the SARC. Following the introduction, for each assessment reviewed, the report should address whether each Term of Reference of the SAW was completed successfully. For each Term of Reference, the SARC Summary Report should state why that Term of Reference was or was not completed successfully.

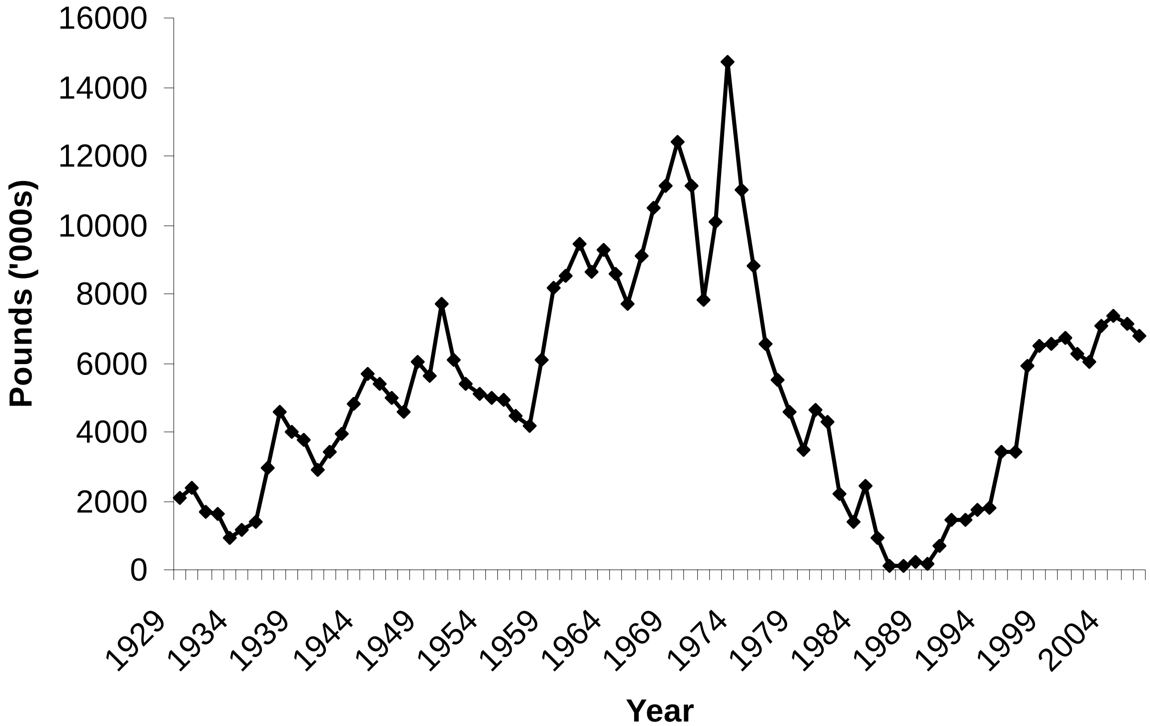
   To make this determination, the SARC chair and CIE reviewers should consider whether the work provides a scientifically credible basis for developing fishery management advice. Scientific criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If the CIE reviewers and SARC chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

   The report may include recommendations on how to improve future assessments.

2. If any existing Biological Reference Point (BRP) proxies are considered inappropriate, include recommendations and justification for alternative proxies. If such alternatives cannot be identified, then indicate that the existing BRPs are the best available at this time.

3. The report shall also include the bibliography of all materials provided during SAW 46, and any papers cited in the SARC Summary Report, along with a copy of the CIE Statement of Work.

   The report shall also include as a separate appendix the Terms of Reference used for SAW 46, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.



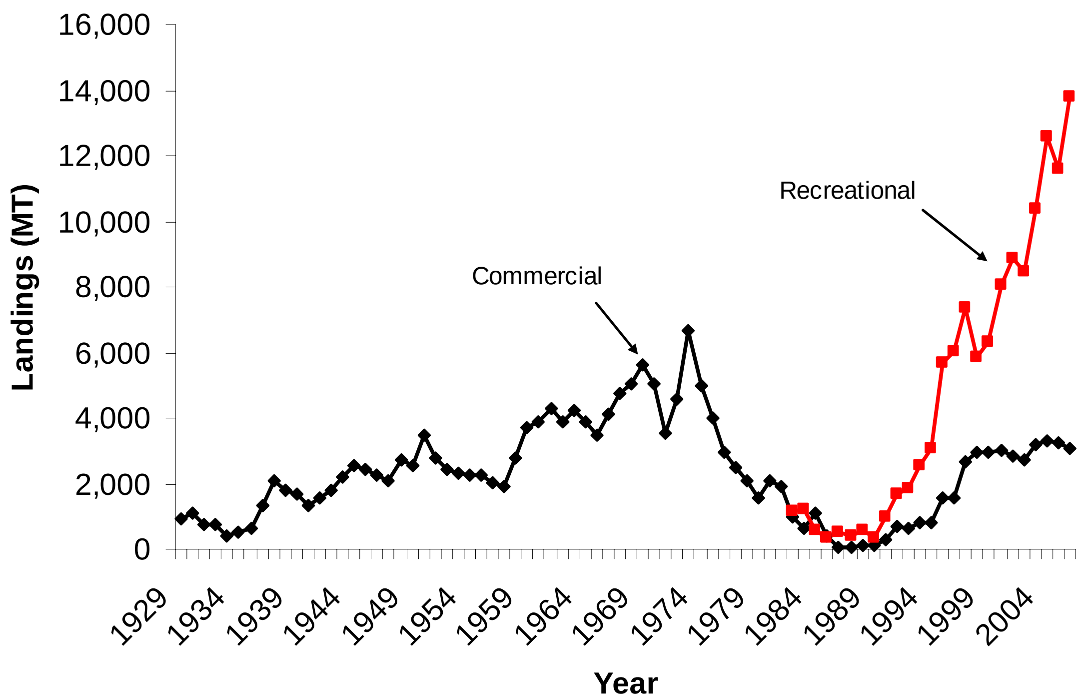


Figure 1. (a) Commercial landings of striped bass (‘000 lbs) along the Atlantic coast of the U.S. during 1929-2006; and (b) commercial landings and recreational harvest of striped bass in metric tons during 1929-2006..

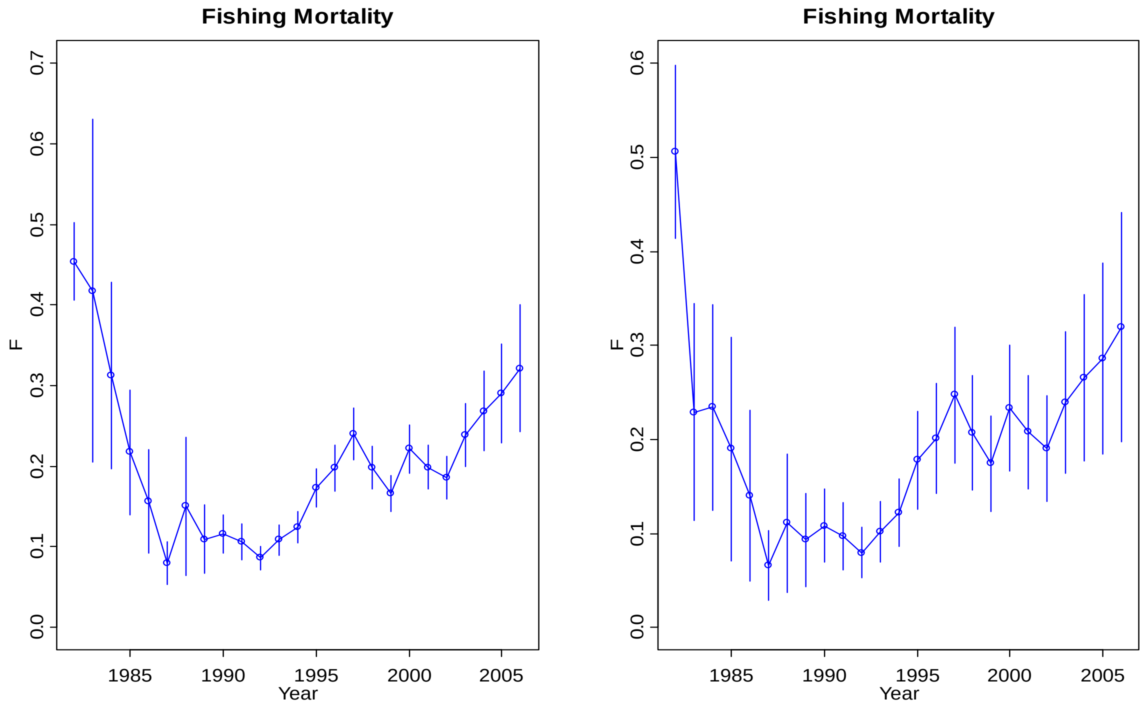


Figure 2. Estimated fully-recruited instantaneous rate of fishing mortality (yr^{-1}) from the striped bass statistical catch-at-age model under the base model effective sample size for the proportion-at-age (left) and under one-tenth the effective sample size (right). Note differences in vertical axis scale when comparing graphs.

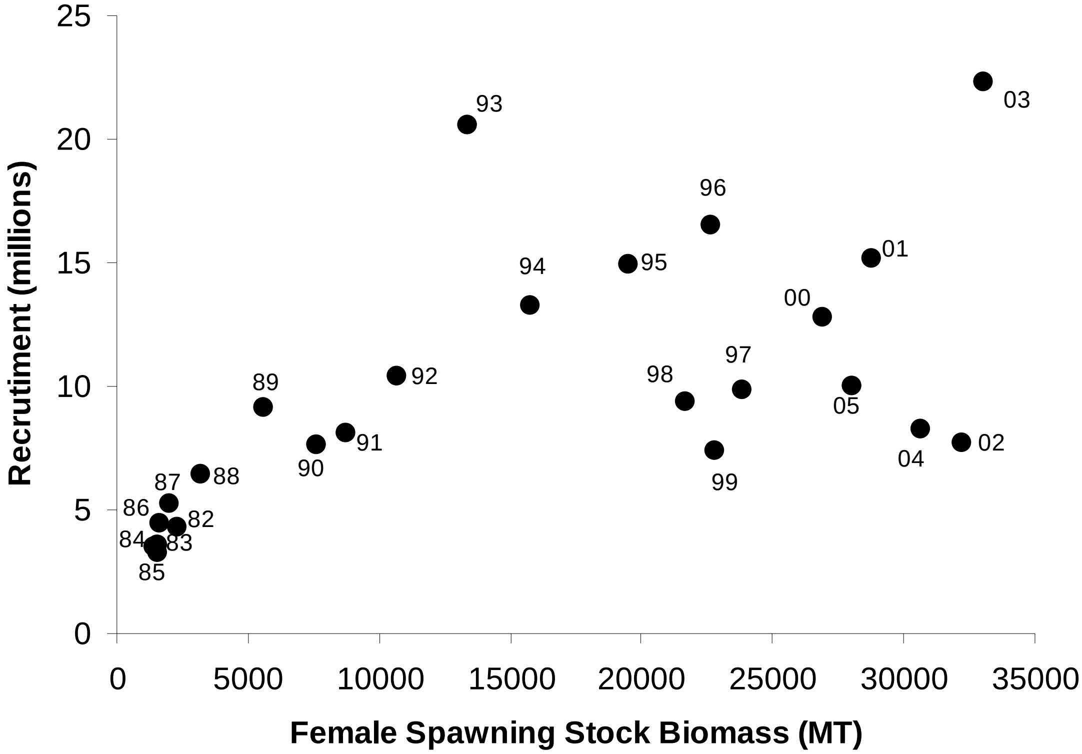


Figure 3. Estimated female spawning stock biomass (metric tons) and subsequent recruitment (numbers of age-1) of striped bass from the statistical catch at age model. The year label represents the year in which the recruits were spawned.